

## 1. 绪论

### 1.1 可视分析

在人类的各类感觉器官中，人眼具有最强的信号处理能力。因此视觉是获取信息的最重要通道，超过 50%的人脑功能用于视觉感知，包括解码可视信息和思考可视符号等。但是，人类的视觉工作内存很小，在认知过程中常常需要依靠外部帮助。图片、图形等在过去两百年间是人类认知过程中最重要的工具。

可视化是人类在认知过程中产生的对事物的认知视图，易于理解并有利于知识交流。而信息可视化是让计算机绘制动态、交互的外部辅助视图，目的是增强人的认知能力。

随着数据规模的增加，直接呈现数据会导致显示空间不足、图元遮挡、绘制效率低等问题。另外，现代分析任务更加复杂，仅仅依赖可视化难以令人发现隐藏在数据背后的模式，迭代式交互分析是领域专家认识和理解数据的重要方法。

为此，可视化专家提出了可视分析概念，它是一种通过交互式可视化界面来辅助用户对大规模复杂数据集进行分析推理的科学和技术。可视分析不仅仅是可视化，它集成了可视化、人的要素（如交互、认知、感知、合作、表示和传播）和数据分析等。相对于可视化，可视分析整合了信息分析、地理空间分析和科学分析等技术，而且人的要素在人与机器的通信过程和决策过程中起到非常重要的作用。可视分析是人的智能（知识和经验）和机器智能（计算）多次协作的迭代过程。

可视分析有效融合了计算机强大的计算能力和人的认知能力，能够发挥二者的优势完成各自擅长的任务。因此，可视分析是交叉融合了可视化、人机交互、认知科学、数据挖掘、信息论、决策理论等研究领域的新研究方向。

可视分析需要合适的载体，而电力系统数据便是合适的载体。电力系统稳定是电力系统中的一个基本问题。电力工业因系统不稳定性会引起严重的停电事故和巨大的经济损失。

电网是电力系统的一种经典形式，被描述为通过发电机向用户传输电力的网络，一个结构良好的分配制度。一个电网是复杂的、动态的、多方面的系统。直到现在还没有一种可行的方法来纪录系统的情况。控制实验对实际的电力系统可能会造成意外的损失。这使得模拟计算性能分析成为唯一可以接受的方法。目前在实际应用中已经验证，模拟计算可以达到比较精确的测试。定期运行数据模拟，对实际事件进行推理，预防性问题分析。

在可视分析中所提供的数据驱动的分析工具，能够有助于分析人员从更多地角度、更深入地探索海量仿真数据中所蕴含的规律，使得电网运行分析具有更强地研究性，适应我国电网高速发展的现状。

### 1.1 电力系统仿真分析

仿真数据，尤其是暂态稳定计算在指导电网稳定运行中起着重要的作用。仿真数据通常是描述每个节点中每个变量的演变的一组时间序列。在模拟它的运行、故障和恢复过程中，不同形式的不稳定性伴随着它的演化过程而出现。数值模拟方法如李亚普诺夫函数及

其变分已被证明是研究动态和瞬态不稳定性的独特手段。因而，分析人员要识别故障事件的发生，研究故障事件的类型和模式，并通过分析变量的时间序列来推断故障事件的原因，是非常麻烦的。

目前缺乏定量描述电力系统稳定状态的理论模型。传统的统计方法，如统计质量控制，提供了基于点的稳定性判断，但对于刻画全局趋势和模糊变化是低效的。

电力系统仿真分析广泛应用于电力系统运行、规划、设计等领域，是电力系统的一项基础性支撑技术。经过数十年的发展，我国基于机理模型的电力系统仿真计算技术已处于世界前列，已有 PSD、PSASP、ADPSS、PS-Model，以及全过程仿真软件等诸多广泛应用的自主知识产权优秀产品，可以基本满足大电网仿真分析的需求。

电力系统仿真分析中的“仿真”是指对电网元件进行数学建模，使用微分-代数方程组表达各个元件的外特性，通过求解这个方程组得到系统的状态变量值，从而掌握整个系统的情况。“分析”则是基于得到的仿真结果进行系统特性研究、控制措施制定等工作。在实际工作中，电网仿真分析可以看作是“提出假设”与“计算验证”的循环过程。

传统的仿真数据分析方式有着各种限制。首先是需要耗费大量人力，电网的日益复杂，使得人力的需求大大增加；其次是十分依赖专家的经验，分析的结论与专家的经验水平息息相关，能否发现问题和能否找出解决方法都取决于专家；再就是容易出现漏洞，由于主要依赖人工分析，难免发生错漏的现象；最后是无法控制误差。

先进的仿真计算技术必须要有先进的数据分析技术和人机交互方式才能充分发挥其作用。虽然目前的各类电网仿真软件均由结果可视化和报表功能，但其着重点仅限于数据的展示，对于数据规律的分析主要还是依靠人工进行只能观察很少的重要元件和断面。这极大地限制了对仿真结果和仿真能力的应用，使得许多在计算结果中蕴含的电网运行规律难以被发现。

基于这些问题，考虑用可视分析来帮助解决这些问题。

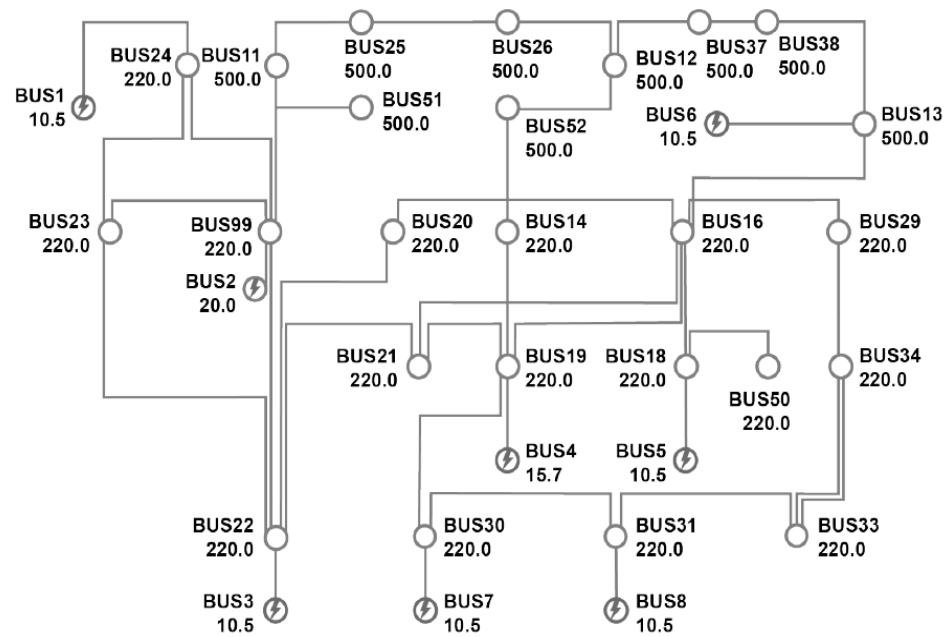


图 1 含有 33 段母线的电网单线图

电力系统的模拟数据包含多个变量及其演化。稳定性状态是所有变量演进的综合结果，要求对演进的细节和表达性视觉证据的可靠表示。

本文基于数值模拟数据，特别是暂态稳定数据，对电网稳定性进行了可视化分析。特别地，将统计质量控制（SQC）方法集成到具有创新视觉设计的稳定性时间的检测和研究中。对基于 SQC 的方法的增强促进了细粒度的稳定性判断，而细粒度的稳定性判断又通过富有表现力的视觉证据得到增强。

将 SQC 方法、可视化编码、领域知识和领域专家的经验无缝结合视为推进智能电网的新机会。我们的工作密切合作领域专家和经验丰富的电力仿真运营商。它们提供领域经验、程序、数据集和分析任务。最终设计和实现了一个综合的视觉系统，以支持多时间序列的视觉探索和分析，识别和区分显著的不稳定模式。

支持电网的异常、不稳定和故障的勘探、识别和推理的可视化分析系统。视觉编码和 SQC 方法的系统集成，为研究可能影响系统稳定性的多个相关因素的联合效应提供可靠的视觉证据。一套用于数字时间序列的视觉设计，使分析人员具有对全球趋势、局部变化和多变量随时间变化的相关性的视觉检查的能力。

## 1.2 电力系统潮流计算

潮流，是电力系统的专业术语，指的是功率。电力系统在运行时，在电源电势激励作用下，电流或功率从电源通过系统各元件流入负荷，分布于电力网各处，称为电力潮流。电力系统是把很多的发电站、变电站、配电站、用户等由输电和配电线路连接起来形成的系统。电力(电能)是在发电站产生的，其中一部分在经过输配电线、变电站和配电站时损失掉，剩下的绝大部分最终被负荷所消耗。这样，从电的产生到被负荷消耗，流过哪一路输配电线，各节点电压是多少，这种计算就叫电力潮流计算或简称潮流计算。

电力系统潮流计算气计算。它的任务是根据给定的运行条件和网路结构确定整个系统的运行状态，如各母线上的电压（幅值及相角）、网络中的功率分布以及功率损耗等。电力系统潮流计算的结果是电力系统稳定计算和故障分析的基础。

潮流分析的目的在于获得电力系统的某个静态运行点，即所谓的潮流断面。由于电网是实时运行的，因此任何对于电网的计算和分析都需要从电网的某个断面出发来进行，例如系统发生故障前一刻的电网运行状态。也就是说潮流计算的结果除了用于潮流分析外，还需要用作其它计算分析输入数据的一部分。

潮流分析可以大体确定电网的一般运行状况（不考虑故障之类的冲击，这种冲击由暂态稳定分析考虑），比如各母线的电压。通过对数据的调整（按照一定要求），可以使潮流计算结果趋于合理，在满足相关要求的同时，也能为其它计算提供良好的基础。例如，如果发现电网内某些母线的电压偏低，可以通过投入无功补偿的方式使之升高。

潮流分析的结果可以用于作为电力系统实际运行的参考，即在实际电网调度中，正常情况下系统的控制边界都在已经计算过的潮流方式之内，比如系统的实际大运行方式不会超出之间计算的潮流大方式。当然，在具体下达调度命令时，调度员还是有一定的自由决定权的，只要系统运行在已经分析过的潮流方式框架之内就可以。

通常情况下，在一次电网分析工作中，需要进行多个运行方式的潮流分析并得到相应的潮流计算结果，诸如系统运行的大方式（重负荷）、小方式（轻负荷）等。基于这些结果，一方面可以确定电网在多个不同运行方式下的潮流分布情况，并采取相应措施；另一方面也能使暂态稳定、小扰动等计算可以基于不同的电网运行方式进行，获得更好地分析效果。

暂态稳定计算在电力系统规划和运行分析中占有重要位置。它不仅为规划系统的电源布局、网络接线、无功补偿和保护配置的合理性提供电力系统暂态稳定性的校核，为制定电力系统运行规程提供可靠的依据，而且可用于研究各种提高暂态稳定的措施并为继电保护和自动装置参数整定提供依据。

暂态稳定计算通常是由一个潮流计算结果开始，以故障作为输入，在某一时刻（每一步迭代）输入到计算中进行迭代，迭代的每一步代表一个时刻。在计算出结果后，研究院通常依靠观察攻角的变化来判断电网系统是否稳定。相比潮流计算，暂态稳定计算输出的物理量的个数更多，因而更有利于我们进行切入。

到目前为止，监控电网的主要方式是通过对暂态稳定计算的结果进行计算和判断，从而掌握电网的运行状态。然而现有的方法仍有许多问题：

一是特征值都是人工提取的，关联关系太多，特征之间的关系不清楚；二是现有的电网监控系统更加注重于特征值的计算与计算结果的展示，未能利用数据可视化的方法将数据统筹到一张图表中，使得监控起来繁杂难用；三是现有的电网监控研究大多都在研究微网，即对城市内配电网的监控与判断，对于大尺度下的高压输电网，即从发电机到城市配电站之间的电网研究少之又少。

## 1.3 相关工作

### 1.3.1 电力系统可视分析

多年来，可视化技术已经在电网数据分析领域得到了应用，为一系列演示和运行需求提供了实用的工具。如图1.2 所示的GreenGrid 系统[1] 和如图1.3 所示的Grid-IE[2] 是最流行的两种工具，可以使用户直观地分析电网及其运行情况，主要用于满足传统数据可视化方法对稳态和静态模拟的基本需求。

Overbye 等人充分利用动画、轮廓、数据聚合和虚拟环境等传统的可视化技术，对功率的传输进行了可视分析[23]。类似地，Guorui Zhang 采用GPS 同步频率测量，实现了如图??的系统，通过在线事件触发和干扰应用的定位来实时监测广域频率状态[3]。

态势感知是分析电网的第一阶段。对此，Pak Chung Wong 和Jun Zhu 均设计了不同的系统来对不同层级的电气配置进行可视化[1;4]，以确定电网的状况和稳定性。其中，Jun Zhu通过利用数据驱动的自动生成算法，成功地将公共信息模型（CIM）转化为了如图1.5 所示的电力网络布局图[4]，为分析电网的拓扑结构做出了贡献。

应急分析是电网分析中的一个重要问题。Pak Chung Wong 提出了针对该问题的可视化分析流水线，他实现的如图1.6 所示的系统可以利用低维散点图表示与地理电网视图，支持用户探索浏览电网系统的细节信息

### 1.1.1 系统稳定性可视分析

稳定性分析在分析一个动态系统中起着核心作用。特别的，可视化可以帮助识别和发现有价值的稳定模式，不同的领域需要各种可视化手段来适应数据特征和领域问题的特定特征。

#### 1.4.2.1 国外研究相关工作

一个由E. Grollier 完成的早期的调查[24] 总结了动力系统稳定可视化的工作，提出了线性和非线性情况下适用的原理和方法（如可视化Lyapunov 指数和相邻的轨迹）[25]。Cokkinides等人应用Lyapunov 间接法，通过计算实时动态变电站的能量函数，并对其计算结果进行了可视化，来观察电网系统的暂态稳定性[26]。Jifeng 等人提出了一种在关键割集空间中的超平面描述的方法，来对大宗电力系统的电压稳定区域（VSR）进行可视化，并且通过包含1305 段母线的华中电力系统验证了该项工作的实用价值[27]。Michael 等人介绍了稳定性存在区域（Region of Stability Existence, ROSE）的概念，并描述了利用相量测量单元（Phasor measurement unit, PMU）数据计算稳定性存在区域和运营边际的框架[6]。其提出的方法是通过精确计算电力系统稳定裕度来实时监测传输系统的自动化过程。分析过程中同时监测电压限制，热限制和稳态稳定性，并将结果域显示在如图1.7 所示的横纵坐标轴分别为不同母线的相角的平面上。Micheal 等人通过使用新英格兰ISO 公司（ISO New England）的实时模型、SCADA[28] 系统和PMU 数据，讨论了人工干预（修复）行为对该地区的影响。研究结果表明，这种方法有效提高了新英格兰ISO 公司的传输网络的可靠性，并可用于防止大面积的停电故障。通常情况下，这些研究采用视觉形式来显示当前操作点与计算稳定性区域或稳定性边界之间的差异。受限于当时可视化技术的发展和算力的缺乏，上述工作多只向用户提供了朴素而单一的视图。这些视图主要用于在直角坐标系上刻画波形、区域和能量函数，一方面缺乏与用户间的交互，另一方面视图与视图之间分割、缺乏联动，并不足以向用户提供用以洞悉全网状态的集成型可视分析系统。

图 2 稳定性存在区域及操作点轨迹

M. D. Flood 最近的工作是对100 多种可视化方法和系统进行了系统性的评估，以监测财务数据的稳定性[29]。其中一个案例是在给定的稳定性定义下，他采用了各项金融数据的聚合值来衡量最终的金融系统的风险[30]。在生物研究领域，Shen 等人通过使用小倍数报告了他们关于DNA 结构的分布和稳定性的研究[31]。与此同时，Alsenaidy 建议使用经验相图（Empirical Phase Diagrams）和雷达图来直观比较蛋白质的物理稳定性[32]。

#### 1.4.2.2 国内研究相关工作

PowerWorld 公司以其丰富的产品线，为传输规划人员，电力营销人员，系统操作人员和培训人员，教育工作者以及任何希望以用户友好的格式访问电力系统信息的分析人员提供了广泛的选择范围。孙惠娟等人通过将IEEE30 节点系统作为仿真对象，证实了性能优越的PowerWorld 软件可为电网稳定性分析提供一个新思路与方法[33]。如图1.8 所示，范国英等人通过将WSCC-3 机9 节点系统作为仿真对象，刻画了各节点电压在电网发生故障时的波动情况；同时，算例结果再次验证了“PowerWorld 软件可以帮助调度人员直观形象地掌握全网运行情况”这一观点

[7]。

图 3 系统故障瞬间各节点电压情况

梁尚捷采用Hadoop 和HBase, 通过Map 函数将暂态稳定计算方法并行化, 加速了该算法的计算速度; 并使用JFreeChart 实时展示了如图1.9 所示的计算结果[8]。一方面受限于实验资源, 该实验并未完全体现出其架构计算能力大的特点; 另一方面, 该实验主要聚焦于通过提高利用算力获得计算结果, 而非致力于可视化技术的探究与实践, 因而缺乏一套高效的人机交互界面。

图 4 朴素的电力系统暂态稳定分析系统首页

上述工作主要聚焦在单变量的分析上, 除去由外国公司开发的PowerWorld 软件, 各项研究工作支持的用户交互手段仍相对匮乏。本文工作与以往的工作不同, 自主研发了一套新颖的可视化分析系统, 来描述可能影响系统稳定性的多个相关因素的综合影响。

### 1.3.3 基于比较的时变数据可视化

随着时间序列的数量和复杂度的增加, 对基于比较的可视分析技术的需求与日俱增。

一般的, 信息可视化的比较设计分为三类[34]:

- 并排法 (Juxtaposition) 通过以时间或空间顺序对时间序列进行了排序[35;36]。如图1.10所示的LeadLines[9] 和如图1.11 所示的CloudLines[10] 将并排法拓展到了大规模数据集上, 实现了在不聚类的前提下细粒度地展示多个时间序列。

图 5 LeadLines 主视图

图 6 CloudLines 主视图

- 交叠法 (Superposition) 将多个对象重叠放置在同一空间位置中, 允许用户用不同的颜色或透明度来直接观察对象之间的差异。如图1.12 所示, Collberg 等人使用了并置法, 将多张研究对象的快照放置在同一空间中以显示图随时间的变化[11]。

- 显式编码法 (Explicit Encoding) 利用针对时间序列数据的不同视觉编码来揭示对象之间的关系或差异[37]。如图1.13 所示的World Lines 是多重集成仿真比较可视化的一个复杂例子, 其视觉设计结合了包括显示编码法的上述三种设计[12;38]。

原则上, 所有的比较可视化技术都可以被运用到电网仿真数据的面向时间的稳定模式的可视化分析中。

图 7 使用并置法刻画图随时间的变化

图 8 World Lines 系统概览

## 1.4 项目研究目标和内容

本项目依照现有的，通过模拟计算所得到的模拟电网特征数据，包括电网各关键节点（我们称之为“母线”）的电压值、频率值，以及模拟电网中的模拟发电机之间的相对功角值，对电网的暂态进行可视化编码与展示。

并且，我们会提供不同的图表格式，包括堆叠图，带状图等，对数据进行展示，并对数据进行统计与特征值提取，计算出一些合理的特征值，展现在项目中，供研究人员进行查阅。

同时，我们可以根据一些控制图计算算法计算出超出阈值的片段，以方便研究员进行后续分析，并可以将研究人员的分析结果记录到后端数据库中，以方便后人查看。

不仅如此，我们还会对大规模数据进行部分适配，以保证即使数据的规模比较大，整个系统的效率也能够合乎要求。

最后，我们会使用真实数据来验证我们设计的所有算法的合理性与准确性，并能给研究人员以提示。

### 1.5 章节安排

本文第二章描述了电网数据及分析任务。第三章制定了本文系统的可视化设计指南，并根据指南和上述分析任务，详细讨论了本文系统的可视化设计方案。第四章通过介绍暂态稳定计算模块、数据预处理模块、后端服务器和前端浏览器，阐述了整个系统的构建与实现。第五、第六章分别给出了用户调研及专家意见。最后，第七章对本文工作进行了总结。

## 数据抽象与任务分析

### 2.1 数据抽象

仿真数据记录了电力网络的运行状态。母线是一种由高导电率的铜（铜排）、铝质材料制成的，供“电站”或“变电站”使用以输送电能的产品。它具有汇集和分配电力的能力。网络中的每一个节点标识了一段母线，而一小部分节点标识了搭载在母线上的发电机。节点之间的连接是连接母线和供电区域的电力线。电气距离是指由母线间电力线路的正序阻和电抗所定义的物理量，指的是电网中成对节点的电气接近度。完善的数学模型可以从电力系统中抽象出节点链路结构，并构建一组包括数以千万计个变量的微分方程。通过改变迭代求解微分方程组时的初始计算配置可以获得一个求值集合，关于这个集合的可视化与分析已经引起了很大的关注。

在本文中，我们研究了用潮流计算模型对暂态稳定数据进行仿真计算。该仿真程序建立了基于电网基础设施（如母线，变压器和负载）的数学建模，将最常见的五类电气故障设置为模型输入，包括母线接地、交流线路断开、发电机退出、线路接地和参数设置错误。仿真结果是一组能够说明电网设备随时间的动态变化状态的时变多变量数据。

通常的，需要计算出多个时变物理量来描述电网设备状态的变化过程。本文作者选择了三个最有意义的特征来支持电网的稳定性分析：电压，频率和相对功角。频率和电压可以用来描述电网中所有节点的运行状态，相对功角则被用来描绘发电机对之间的状态。如果电网运行稳定，没有发生电气故障，功角值将保持不变；这个值的变化和波动只有在触发故障时才会发生。例如，当电气故障发生后，若电压值保持波动且不能收敛到一个恒定值，则电网



很可能不稳定。类似地，这些特征值的任何异常变化都可以被认为是系统不稳定的暗示。由于电网的互联互通，特征的波动可能随时间而相互影响，使得电网稳定性分析更具挑战性。

总之，仿真数据记录了电气故障前后一系列物理特征的计算结果。采用传统的统计分析技术和折线图等可视化图表，电力分析人员可以观察到故障的发生，但很难推断故障事件的具体类型和模式，而如何描述不同变量间的时变关系也就变成了一项更具挑战的任务了。因此，本文合作者希望可以以一种直观客观的方式来刻画故障事件中不同变量随时间变化的相互影响。他们认为，这些特征的数值和统计值对于稳定性分析是有价值的，因为它们与整个系统的能量密度和分布有关，发明直观而全面的稳定性可视化分析方法具有重大意义。

## 2.2 任务分析

本文合作者从一开始就参与了这项研究，他们都对于如何对任务进行分析有丰富的经验。笔者也会与合作者们进行定期的讨论，并从讨论中相关的要求和各个方面的改进。每次讨论都会有对于电力系统分析有着丰富经验的领域专家在场。分析任务和要求在每次讨论中不断地更新改进，用来帮助系统进一步的设计。总体来说，本系统承担了以下几个任务：

判断电网是否稳定。电网的稳定性分析是非常重要的一层，也是电力系统分析上的专家们十分关注的问题。电气故障一旦出现，电网有可能会发生无法逆转的变化而进入不稳定的状态，或者会在经过波动后回复到稳定的状态。电网是不是能在故障触发后保持稳定是一个很重要的标准。

任务二，回答在何种情况下电网可能会进入不稳定状态。不稳定的电网可能会有不同的表现形式，即不同类型的不稳定性在时变多变量数据集上可能会表现出不同的特征。分析师期望找出是何种故障（母线接地、交流线路断开或发电机退出等）导致了当前的电网不稳定状态。

能够辨别电网在何种进入不稳定的状态。不稳定性有着不同的类型，而这些不同的类型在时变多变量数据集上会展现出不一样的特征，这些就使得不稳定的电网展现出不同的表现形式。合作者的期望是找出什么类型的故障导致了电网不稳定状态的出现。

任务三，找出电网不稳定状态下的典型模式。在确定故障类型后，分析师希望进一步地调查电网失稳状态下的典型特征（如周期性失稳或单摆失稳）。这些特征不仅可以用来对电网进行不稳定类型分类，同时可以帮助分析师更好地理解电网的失稳状态。

找出典型的电网不稳定状态。这是在确定故障的类型后，合作者们期望可以对此找出电网不稳定状态下的经典特征（如周期性失摆或者单摆失稳）。这些特征都可以用来对电网进行不稳定类型分类，也可以来帮助分析时了解电网的不稳定状态。

任务四，找出导致电网失稳的具体原因，如故障出现在具体哪段母线上等。本文合作者对电网不稳定的情况推理表了现出强烈的兴趣。了解电网变得不稳定的原因不仅有助于采取相应的措施，从而减轻电力故障带来的成本，也有助于防止这种问题再次发生。

找出导致电网出现不稳定的具体原因。例如故障出现在具体哪个部分等。文章合作者更希望可以针对电网的不稳定进行推理。这些有助于针对不稳定的原因采取合适的措施，进而



减轻电力故障造成的损失，也有助于防止这种问题再度出现。

为完成上述任务，本文作者与合作者讨论，总结出了六个要求：

要求一：描述整个电网的拓扑结构。应该给出电网及其配置的概述，以支持对总体结构和模式的认知。

能对整个电网进行描述的拓扑结构。给出电网及其配置的大体总览，方便对于总体结构和模式的认知。

要求二：刻画整个电网的运行状态。计算出的时变多变量数据预示着全网状态的演变，一个合理的交互界面有助于展示并鉴别全网变化趋势及每个变量变化的细节。

对于整个电网的运行状态进行刻画。计算出的时变多变量数据代表了电网状态的演变，需要一个合理的交互界面帮助鉴别和展示电网数据的变化趋势，以及变量变化的相应细节。

要求三：异常检查。异常的值（某一时刻的值）或模式（某个时期内值的变化）很可能表征了电网的不稳定性。因此，自动异常值检测算法对于支持异常值的有效识别是必需的。检测到的异常应该通过独特的视觉编码自动高亮，以便分析人员能够根据其领域的专业知识和经验发现异常模式。

异常的检查。异常值的变化代表了电网的不稳定性。因此必须要有一个有效识别异常值的自动异常检测算法。

检测到的异常通过独特的视觉编码提供高亮，以便于分析人员可以根据经验和相关领域的知识发现异常模式。

要求四：检测多个变量的时序数据。不同变量的值会随着时间的推移而变化，这点在发生大型故障时尤其明显。系统应该支持在某个时间点或某个时间段内对变量进行逐一观察，以便分析人员能够专注于更详细的特征变化。而且，视觉编码和设计应该能够说明不同变量的周期性变化和异常模式。

对于多变量的时序数据的检测。不同变量的值会随着时间的推移而发生变化。在发生大型故障时最为明显。系统需要支持在某个时间点或者时间段对于变量进行逐个观察，使得分析人员可以更专注与详细的特征变化。而且，视觉编码以及设计需要能够表现出不同变量的异常模式和周期性变化。

要求五：设计成对时变序列的可视化方法以提供观察“时序多变量数据对异常演变的影响”的直观手段。除了研究个体变量的变化之外，对变量对（如电压和频率）的比较和匹配检测对于不稳定性的分析具有重要价值。

设计出对时变序列进行可视化的方法，使得可以直观的观察“时序多变量数据对异常演变的影响。”除了研究单个变量的变化以外，对变量对（例如电压与频率）的比较以及匹配检测也有很重要的价值。

要求六：显示电网存在的局部扰动。一旦电网波动后稳定，大部分特征值趋于保持不变。然而，在整个电网保持稳定的情况下，多个节点的频率扰动仍然存在。因此，除了揭示不同变量的整体变化外，还应展示电网内存在的局部扰动。

对于电网存在的局部扰动进行展示，一旦电网波动稳定后，大部分的特征值趋于不变。但是同时，多个节点的频率扰动仍然还在。因此，除了展示不同变量的整体上的变化，还要展现电网内的局部扰动。

### 2.3 本章小结

本章先讨论了电网系统的数据抽象及输入：将母线及电路线抽象成一张无向图，并以五种不同的电器故障作为模型的输入；再描述了本文系统应承担的四个任务：判断系统是否稳定，回答在何种情况下电网可能会进入不稳定状态，找出电网不稳定状态下的典型模式，以及找出导致电网失稳的具体原因（如故障出现在具体哪段母线上等）；最后，总结了本文系统欲完成上述任务所需满足的要求。

## 第 3 章可视化设计

### 3.1 可视化设计目标

在研发之前，系统设计思路要根据领域专家的意见反复评估和提炼。本文作者提出以下四条可视化设计指南，使得本文的可视设计可以适配特定的领域特征：

1. 针对特定任务的解决方案需要对不同对象和数据样本进行可视化分析。而且，各方面的分析都应纳入设计过程，如定量分析和比较分析。现有的用于稳定性分析的可视化尝试大多是分散的，因为单个问题的解决方案总是被分解成一系列独立的视觉形式。相反，本文工作应寻求一个系统的设计方案，实现所有需要的分析方面的综合。最终设计应是所有研究点紧密相关的观点的组合，从而形成一个完整的系统的可视化解决方案。
2. 可视化设计应该容易理解。我们大多数的合作者缺乏可视分析的经验，他们唯一的尝试是曾经对时间序列进行过检测。因此，易于理解的直观可视化设计将有利于对数据进行探索，从而最大限度地降低学习成本。比如，可以考虑自然的方式结合基本的统计图表和基于时间线的设计等。
3. 视觉设计应该是物理上有意义的。模拟数据记录物理系统的动态，因此数据处理方法和视觉编码都应该保持物理意义。
4. 用户认知和交互应该是直观的。视觉认知和交互使得在知识发现过程中涉及领域专业知识成为可能。领域专家通常不熟悉复杂的交互，因此，不应提供相对复杂的视觉编码和交互。

### 3.2 可视化设计

本系统的核心在于使用统计和可视化方法来检测和研究电网系统的不稳定性。图 3.1 展示了系统概览。本文作者受起伏的潮水的形状所启发，提出了一种用来编码时变多变量数据和与暂态稳定密切相关的模式的可视化设计（3.2.1 和 3.2.2 节）。同时，本文系统中所含有的辅助视图（3.2.3 节）和用户交互界面（3.2.4 节）进一步的增强了对稳定性相关特征

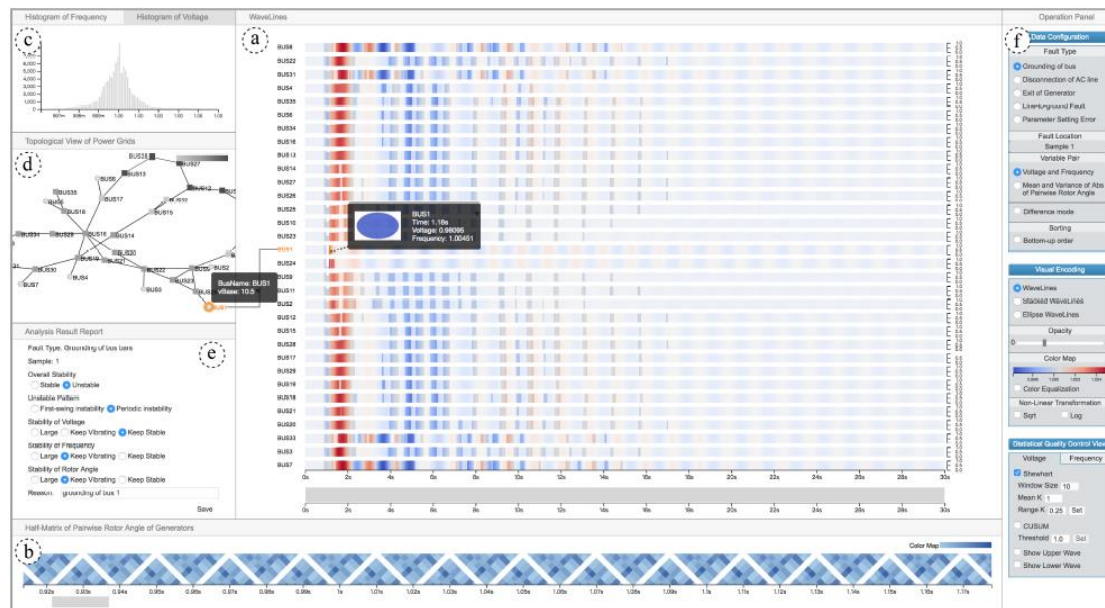


图 9 WaveLines 系统概览

### 3.2.1 基于统计质量控制算法的可视化设计

#### 3.2.1.1 累积和与休哈特控制图算法

统计质量控制方法 (Statistical quality control, SQC) 被广泛用于监测和控制一个过程，目的是确保合格比例的确认产品。在各种 SQC 方法中，控制图方法在测量过程稳定性和检测显著变化方面表现出色。SQC 控制图理论构成了我们异常检测方法的基础。

具体而言，两种控制图很好地符合我们的要求：累积和控制图 (CUSUM chart) 和休哈特图 (Shewhart chart)。两者都能够从变量的时间序列中检测异常间隔。在累积和控制图方法中，如果时间点的累计和与平均值之间的差值高于给定的阈值，该时间点则会被标记为异常[9]。所有检测到的时间点一同组成了累积和图表。算法 1 描述了适用于电网模拟数据的累积和算法。

休哈特方法试图计算平均值和极差 (range) 从一个标准值到均匀分配区间的移动，并确定具有较大偏差的值，从而产生两个以平均值和极差表示的离散时间间隔序列。针对本文所处理的时变数据，假设某一物理量的时变序列  $X$ ,  $x_i$  表示该物理量在时刻  $i$  的值 ( $i=0; \dots; n-1$ )，对于时刻  $i$  该物理量的采样，样本容量为 1，亦即样本中只有一个值  $x_i$ 。若直接使用该采样，平凡的，该时刻的平均值和极差为  $x_i$  和 0，不能很好地发挥休哈特方法检测较大的漂移 (shift) 和异常点 (outlier) 的能力[41]。

为了逐步分析随时间的稳定性，本文作者提出了通过滑动窗格来进行时间分割的方案：将与时刻  $i$  相近的时刻一同作为时刻  $i$  的采样  $W_i = [x_{i-s}; x_{i-s+1}; \dots; x_i; \dots; x_{i+s-1}; x_{i+s}]$ 。算法 2 描述了本文作者对休哈特算法的改进细节。

本文作者与在合作者进行了多次实验，以验证休哈特算法的有效性后，选取合适的阈值  $t_1$  和  $t_2$ 。对于某一正常时刻对应的滑动窗口  $W_i$ ，其对应的平均值  $m_i$  与故障发生前的正常值的差应小于阈值  $t_1$ ，亦即  $|m_i - \mu| < t_1$ ，其中  $\mu$  是时变序列  $X$  的标准偏差。这个设置在统计上是有意义的，因为在正态分布中，一个值超出该值的概率小于 0.3%。同样的，极差的阈值  $t_2$  被设置为  $0.25R$ ，其中  $R$  是整个时变序列  $X$  的极差。若  $|m_i - \mu| > t_1$  或  $R_i > t_2$ ，则时刻  $i$  将被标记为异常。

电网遭受到线路接地故障 (Line-to-ground Fault) 时会展现出不同的状态：失去稳定或保

持稳定。表 3.1 展示了针对不同全网状态，采用休哈特方法和累积和方法来检测得到的可视化结果。其中，全网状态由本文合作者在对两个不同的算例进行分析后标记得出，累积和方法的阈值  $t = 120$ ，休哈特方法的参数为  $s = 2$ ;  $t_1 = 3$ ;  $t_2 = 0.25$ 。可以发现，当电网受到冲击后恢复稳定时，休哈特方法显示全网不存在异常时刻，而累积和方法存在着“误报”的情况；当电网受到故障影响并最终失去稳定时，休哈特方法仍能检测出全网中存在的异常时刻，同样的，累积和方法也根据其自身的检验标准，给出了参考的异常时刻。值得一提的是，使用同一组参数对不同算例进行休哈特方法检测时，检测结果相对稳定（不需要频繁地修改参数）。但使用累积和方法时，往往需要大范围地对参数进行修改。

在一定程度上，累积和与休哈特方法都满足检测时间序列异常的要求，但对比上述检测结果，出于以下三个原因，本文作者所在团队认为休哈特方法比累积和方法好：首先，累积和方法在假设样本遵循正态分布的情况下检测平均值的变化，这在大多数情况下不成立。

其次，当发生故障时，基于当前值与全时刻平均值的偏移来判断当前时刻电网是否存在异常可能失败，因为平均值本身可能在很大程度上偏离了电网在稳定状态下的平均值。第三，因为累积和方法会考虑历史（或近期）数据对当前时刻的影响，采用这种方法方法可能影响对电网从不稳定状态恢复到最终稳定状态的判断。

#### 3.2.1.2 基于休哈特特征的可视设计

本文作者试图对两个变量的时间序列（如电压与频率）检测到的平均值和极差的异常值进行编码。这么做有可以提供一种直观手段来观察“时序多变量数据对异常演变的影响”（要求五）。

如图 3.2（a）所示，对于任一时刻  $i$  对应的椭圆形图标，它的编码方式如下：竖半轴的长度编码了电压的均值；横半轴的长度编码了频率的均值；在逆时针方向上定义域为  $[0;_4]$  的倾角编码了电压的极差；背景色用从蓝到红的值域编码了电压的极差。

这种图标可以表征在某个时间点两个变量的基于休哈特控制图算法的局部特征。当研究变量对的基于休哈特特征的全局变化趋势时，在时间线上显示一系列基于休哈特控制图算法的图标也是非常有用的（图 3.2（b））。除了上述的可视设计之外，本系统不仅可以对算法 1 与算法 2 的检测到的异常区间进行高亮，还支持关于不同区间集合的交运算和并运算。

#### 3.2.2 WaveLines 的可视化设计

WaveLines 旨在监视整体状态，描绘全局模式，并分析局部变化。为了满足设计要求（要求二至六），本文作者仔细选择了传统的彩条作为基本的视觉形式，并提出了两种视图：WaveLines 和堆叠式 WaveLines。前者将每个色条分别沿着时间线放置，而后者则将所有色条从下到上堆叠起来，形成一个全局的流形图形状。

上述两种可视化方法都使用水平彩条来编码时间序列。常规解决方案使用颜色对一个变量的时间序列进行编码，而我们在每个时间点采用条带的颜色和高度来编码两个变量，从而观察成对变量（要求五）随时间变化发生的波动。

遵循设计指南三，本文作者根据电压、频率和发电机间相对功角的物理意义构造了两种类型的变量对。第一种变量对由电压和频率组成，因为它们是表征电网中所有母线的物理状态的两个最普适的变量。因此，母线的两个时间序列（电压和频率）可以分别用任一视觉通道（高度或颜色）来表示。其次，相对功角可以用来测量两个发电机（电网拓扑中的特殊节

点)之间的相对差异。一个发电机与其他发电机之间的相对功角组的平均值和方差可以表示从发电机到其他发电机的基于功率的电气距离。因此,一个发电机与其他发电机之间的相对功角组的平均值和方差的两个时间序列同样可以上述视觉通道来编码。为了简单起见,我们通过在本文的其余部分应用电压-频率对来说明 WaveLines 和堆叠式 WaveLines 的作用。

### 3.2.2.1 WaveLines

图 3.3 给出了一个例子:每条母线对应的条带垂直排列,每条条带用其长度和颜色分别编码母线电压和频率的时序数据。本文作者选用了由 K. Moreland 精心计算、设计的蓝色到红色的颜色过渡来描述频率的变化[42;43]。由于 WaveLines 很好地保持了动态系统的时间一致性,所以其既可以用来表示这些物理量平滑过渡,又可以用来刻画它们随时间推移而发生的剧烈变化。WaveLines 具有如下特点:

- 波形和颜色变化为证明电网稳定性提供了直观的证据,并呈现其他特征,如特征变量(要求二、四)波动幅度和频率等。
- 波形与颜色之间的匹配程度意味着两个变量的共现模式,进一步的,可以体现出这两个变量不同的演变(要求五)。
- 可以通过比较两个波浪的形状和颜色来检查两段母线或发电机(要求四)之间的差异。
- 波形与颜色沿时间线的大范围波动可能表示电网中存在电气故障(要求三)

### 3.2.2.2 堆叠式 WaveLines

WaveLines 表示保留了的时间相关性,但是由于条带是分开布置的,因此不能有效地呈现不同母线之间可能存在的一致性。因此,本文作者提出了堆叠式 WaveLines 表示法很好地弥补了这个问题(图 3.4)。堆叠式 WaveLines 表示法有一个特有的优点:它可以揭示电网的一小组节点间存在的局部扰动(任务二、六)。当某些变量(通常是频率)发生局部波动时,通过观察堆叠的波形中的颜色变化模式,可以快速定位可能的局部扰动。

### 3.2.2.3 对时变数据进行排序

无论是波纹图还是堆叠式波纹图,将母线沿纵轴排列对于揭示它们间存在的显著模式都是有幫助的。排序操作可以分三步进行:

1. 构造电网的有向无环图:其中每个节点代表一段母线,每条边表示相邻节点的电气距离。
2. 计算最短路:在计算任一节点离故障中心(可能是多个故障中心)的最短距离时,本文作者通过预处理,将多个故障中心原本用于解决单源最短路的 Dijkstra 算法计算故障节点(可能有多个故障节点)和其他节点之间的最短路径。
3. 对母线进行排序:本文作者根据节点链接图中与故障中心的最短路径长度,采用了两种不同的排序方法。第一种是将故障点放置在垂直轴的中心,并将其他节点按照它们与故障节点的最短路径长度向上或向下放置。第二种将故障点放置在底部,并且将其他节点按照它们与故障点的最短路径长度依次放置,最短路径距离越小的节点离故障点放置得越近。在案例调研章节中,读者可以发现两种排序方法均揭示了故障传播的影响与电气距离成正比的现象。

### 3.2.2.4 设计细节

直观的交互手段有助于用户认知(要求四),本小节描述了关于 WaveLines 的设计细节。

#### 差分时间序列

为强调变量相较正常值(通常表示为发生故障前的值)发生的偏离,本文作者通过计算变量的正常值和时间序列之间的差值,得到了变量的差分时间序列。这为 WaveLines 表示提

供了一个额外的模式，称为差分模式。差分模式放大了变量随时间的变化，因此非常适合描绘异常演变。

图 3.5 示出了两个电压和频率差分时间序列的 WaveLines 图，其中波形和从红色到蓝色的周期性颜色变化表明电压和频率的快速波动。另外，电压和频率的变化模式不一致：一个正在增加而另一个正在减小。由于偏差可能是负的，所以堆叠式 WaveLines 不能用于显示差分时间序列。

### 均衡化

考虑到变量的值的分布可能是不均匀的，均衡化方法可以使原本不均匀分布的值分布均匀。本系统支持用户交互式地启用或关闭均衡化操作，算法 3 描述了本文作者设计的均衡化算法。

### 对数值的非线性变换

当发生故障时，电压和频率等变量剧烈波动，导致较大的数值范围。但是，高度和颜色的分辨率都很低。均匀地将大范围的值映射到低分辨率的可视通道中会抑制那些值相对较低的项（时刻）。本文作者的解决方案是通过对变量值应用平方根或对数运算符来进行非线性变化，重新缩放范围。

### 总量曲线

如图 3.6 所示，本系统可以为选定的变量，如电压，绘制沿 y 轴变量和的辅助曲线，作为添加在 WaveLines 或堆叠式 WaveLines 视图上的背景，用来显示该变量在整个电网中的总体变化趋势。

## 3.2.3 辅助视图

### 电网拓扑视图

拓扑视图（图 3.1（d））显示了整个电网的拓扑结构（要求一）。本文作者使用节点链接图表示来描述图结构。如果输入数据包含位置信息，则节点的位置是指母线的地理位置；反之，通过标准的力导向图布局算法来放置节点。每个节点的形状表示是否存在发电机：具有闪电图标的圆形表示该段母线上存在发电机，方形表示该段母线上不存在发电机。节点颜色编码该段母线的标准电压。另外，发生故障的位置附近的节点用橙色的圆环高亮显示。

在用户对 WaveLines 或堆叠式 WaveLines 视图进行探索时，此拓扑图将高亮出对应的母线和其邻边。

### 半功角矩阵视图

本文作者采用半功角矩阵视图（图 3.1（b））来描述发电机中成对相对功角随时间的演变（任务四），其中半矩阵的每个单元格的颜色编码了相对功角的大小：颜色深的代表相对功角的绝对值较大，颜色浅的代表相对功角的绝对值较小。因为相对功角的值是对称的，发电机  $i$  对发电机  $j$  的相对功角满足： $RA_{ij} = RA_i \ominus RA_j = \ominus RA_{ji}$ ，使用半矩阵视图可以节约空间，减少用户眼球的移动距离，提升其工作效率。

### 统计直方图视图

统计直方图视图（图 3.1（c））提供了对值的分布的概览（任务二）。

### 诊断视图

如图 3.1（e）所示，用户在与系统交互的过程中，可以利用诊断视图记录发现的特征，并生成诊断报告。该诊断报告记录了电网在给定采样下的整体稳定情况，以及电压、频率和相对功角的稳定特征。当用户重新打开该样本时，可以查询到此前保存的诊断报告。

## 3.2.4 交互设计

鉴于分析师在并不具备丰富的可视分析经验，本文系统为其提供了简单直接的交互操作（设计指导四）。

### 突出高亮

如图 3.7 所示，当某个区间被统计质量控制方法检测为异常区间或鼠标悬停在某条条带上时，该区间或条带会被高亮。

### 按需显示

类似于传统的可视化系统，本系统支持用户通过在视图或时间轴上的自由缩放来浏览（堆叠的）WaveLines 视图（效果如图 3.8 所示）。

### 动态链接

如图 3.9 所示，本系统支持多个视图之间的动态链接：当鼠标悬停在拓扑视图中的某个节点时，该节点和 WaveLines 视图中相应的条带将自动高亮。

## v3.2.5 大规模电网可视化设计

一个真正的电网可能包含数以千计的母线。为了证明本章的可视化设计是可扩展的，本文作者通过逐层求精的交互来设计和实现了一个简单的原型。

本文作者首先通过对所有母线进行聚类来预先计算分层结构。图 3.10 展示了大规模电网下的 WaveLines 可视化效果，该视图沿垂直轴分为三部分：中间部分用于提供所选母线的细节，而另外两部分用于聚合上下文，即其他母线。中间部分的最大容量是预定义的，例如 20 段母线，在该区域内的母线用标准的 WaveLines 展示。其他母线的时序序列首先在所述的层级中被聚合（值平均值），并且用 WaveLines 来显示。分析师可以通过移动视图左侧的定位滑块来交互式地更改显示的细节（中间部分）及其他被聚合的内容（上下文）。算法 4 给出了对上下文母线进行聚合的具体实现。

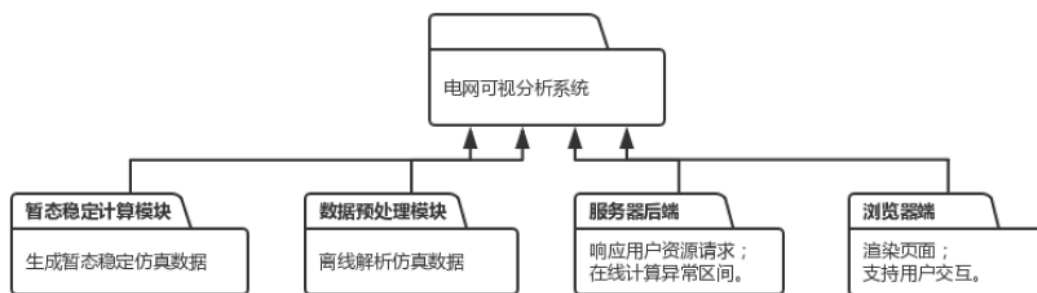
## 3.3 本章小结

本章首先阐述了进行适配电网系统可视化设计过程中应遵循的可视化设计指南，接着，介绍了本文的两大贡献点：基于统计质量控制的可视化与 WaveLines 可视化，随后，介绍了本文系统中所含有的辅助视图及支持的交互设计；最后，不失一般性地讨论了针对大尺度电网系统的可视化设计。

## 系统设计 with 实现

本系统采用浏览器-服务器（B/S, browser and server）架构。B/S 架构与客户端-服务端（C/S, client and server）架构不同，它具有相对平台独立、易用和易开发的特点，这些特点分别体现于：B/S 架构的终端运行在浏览器上，软件运行的最终效果与操作系统较不相关；软件的使用成本低，用户只需处在能对服务器发起会话的网络中借助浏览器即可访问系统，无须引入可能更为复杂的环境配置等环节；JavaScript, Python 是目前比较流行的语言，它们背后强大的社区能为使用这门语言的开发者提供迅捷有效的支持，可以更为快捷的实现想要的效果。





依据数据通路和各模块间的依赖关系，可将本文系统划分为四个模块：暂态稳定计算模块、数据预处理模块、后端服务器和前端浏览器。本章会对这些模块的具体设计进行介绍。

## 4.1 暂态稳定计算模块

暂态稳定计算模块可以生成电网仿真数据，该模块产生的数据包括但不限于：记录每个节点（母线）静态属性的潮流母线数据文件、记录每条边（基准电压不同的母线间的连线）静态属性的交流线和变压器数据文件、定义故障表现方式（故障类型、出现的时刻等）的网络故障数据文件、定义结果文件格式的暂态稳定计算结果元数据文件以及暂态稳定计算结果文件。

## 4.2 数据预处理模块

数据预处理模块针对原始仿真数据中存在的数值重复、数据缺失等问题，首先对数据进行清洗，再从潮流母线、交流线和变压器数据文件中提取出电网拓扑图。Dijkstra 算法常用于解决不含负权环的单源最短路问题，本文作者实现了变种的 Dijkstra 算法<sup>5</sup>用于解决电网系统中多故障源的“多源”最短路问题。进一步的，本模块根据暂态稳定计算结果元数据文件，解析了暂态稳定计算结果文件，为不同的物理量生成了不同的 Json 文件，存储在服务器端以供用户按需访问。最后，根据网络故障数据文件，本模块可以计算故障发生后每段母线上任一物理量与故障前一时刻的差分，这些差值所形成的差分文件将在生成差分 WaveLines 图时发挥作用。

当两段母线间，存在多条线路直接或间接地将它们相连时，它们间的并联组框可以通过如下公式算得，或者近似的，可以用多条道路中阻抗最小的一条支路阻抗作为近似值。

## 4.3 后端服务器

后端服务器负责用户验证与授权、-应用户的资源请求，并根据用户指定的统计质量控制方法及给出的参数，在线计算异常区间。后端服务器是由 Django 框架实现的。Django 是一款免费的、开源的高级 Python Web 框架。它由一批经验丰富的开发者构建，能够代替它的使用者处理许多在 Web 开发过程中遇到的麻烦，使得开发者可以专注于编写应用程序本身，快速迭代产品。

### 4.3.1 用户验证

用户验证的目的是在于验证用户是否确实是其所生成的角色，用户授权的目的在于定义验证用户的资源访问权限。为了简单起见，我们在后文中将用用户验证去指代这两种任务。一个用户验证系统应包含：用户、权限（一个用于标识特定用户能否访问某种特定资源的标识）、用户组、用于登录或限制访问内容的表单和其他视图工具以及一个可定制的密码哈希

系统。Django 中的验证系统采用了非常通用的设计，有些 Web 验证系统中常见的有些特性并没有被其支持。相反的，这些常见问题的解决方案可以从第三方库中获得。这些常见问题有：密码强度检查（检查密码长度、是否含有不同类型的字符等）、最大试错次数和基于第三方平台的验证方式等。

#### 4.3.2 RESTful 资源请求架构

当下是一个各种智能终端（如手机、平板、桌面电脑和其他专用设备）如雨后春笋般涌现的时代，一套规范的应用程序接口（Application Program Interface, AP）命名方式可以很好地帮助前后端通信的设计与实现，使得应用开发者可以快速、正确地开发程序，并提高程序的可维护性。Fielding 提出表现层状态转换（Representational State Transfer, REST）原则，定义了超文本传输协议（Hyper Text Transfer Protocol）请求中的资源（Resources）、表现层（Representation）和状态转换（State Transfer）。本文系统的 API 设计具有结构清晰、符合标准、易于理解、扩展方便的特点，是一个可以很好地响应用户（前端）请求的 RESTful 架构系统。

#### 4.3.3 缓存算法

当后端服务器接收到用户请求时，不失一般性的，后端服务器将通过磁盘 IO 访问特定的样本数据。在计算中，高速缓存（Cache）是存储数据的硬件或软件组件，因此可以更快地为将来的数据请求提供服务；存储在缓存中的数据可能是早期计算的结果，或者存储在别处的数据的副本。当缓存满了时，算法必须选择丢弃哪些条目（Entry）来为新的缓存留出空间。高速缓存算法（Cache Algorithm，也经常称为高速缓存替换算法或高速缓存替换策略）可以用来优化计算机指令或算法，使得计算机中的硬件结构或程序能更好地管理、使用高速缓存。

为保证可以高效、低延迟地响应用户请求，本文作者在算法 6 中实现了一个基于最近最少使用策略（Least Recently Used, LRU）的缓存管理功能来减少系统对磁盘的访问次数。该实现的特点在于调用 get 和 put 函数的时间复杂度都是  $O(1)$  的。这是由于，在给定迭代器指定插入删除的位置时，链表的插入和删除的时间复杂度都为  $O(1)$ ；访问通过哈希算法实现的映射表的时间复杂度同样是  $O(1)$  的。

#### 4.3.4 大规模电网母线聚合算法的优化

假设母线数量为  $M$ ，时刻数量为  $T$ ，算法 4 中 Agglter 函数的时间复杂度为  $O(MT)$ 。考虑到用户在探索大规模电网的过程中会频繁地拖动左侧的滚动条，每一次拖动前端都将向后端服务器发送 AJAX 请求[48]。倘若后端服务器每收到这么一次请求都需要耗费  $O(MT)$  的时间复杂度来调用 Agglter 无疑是十分耗时的。为此，本文作者设计了一种空间换时间的策略，以  $O(MT)$  的空间复杂度和  $O(MT)$  时间复杂度的预计算，将每次后端服务器处理 AJAX 请求的时间复杂度由  $O(MT)$  变为了  $O(T)$ 。

### 4.4 前端模块

#### 4.4.1 模块构建工具

前端模块负责快速绘制可视化结果，提供交互界面等任务。本模块引入了 Webpack[49] 和 Babel[50]，在程序开发和程序发布两个环节都为软件开发者提供了极大的帮助。Webpack 工具不仅支持通过 dev-server 监听 HTML 源文件的改动对网页进行热更新，还可以将多个 JavaScript 文件打包成一个 JavaScript 文件。

Babel 作为 JavaScript 编译器，紧跟时代潮流，支持软件开发者使用 ECMAScript 2017(ES8)

语法。上述特性不仅可以帮助软件开发者更快地开发和发布产品，同时发布产品时仅需向外界提供被 Babel 解释过的代码，很好地保护了源代码。

在开发现代软件的过程中，开源社区不仅使用 ES8 新标准，还采用了一系列拓展的 JavaScript 语法来加强程序语言的表达能力。为适应这一潮流，本模块还采用了如 babelplugin-transform-object-rest-spread 等的 Babel 相关插件，用以支持如对象解构赋值等的高级语言特性。

#### 4.4.2 交互组件库

为实现具有丰富交互特性的前端模块，本文系统采用了 React, D3 等前端交互组件库，支持用户在 SVG 或 canvas 元素上完成探索任务。

##### 4.4.2.1 React

React 由 Facebook, Instagram 及由个人开发者所组成的社区所维护，是一种用于构建用户交互界面的 JavaScript 第三方库。它具有如下的特点：

- 声明式：使用 React 创建交互式用户界面十分轻松，它能为应用程序中的每个状态设计简单的视图，并将在数据更改时有效地更新和渲染正确的组件。声明式编程可以使得代码更具可预测性，更易于调试；
- 组件化：用户可以构建具有自我状态管理能力的封装组件，然后将其组合成复杂的用户交互界面。由于组件逻辑是用 JavaScript 而非模板编写，因此可以通过应用程序轻松传递丰富的数据，并将状态保留在文档对象模型（Document Object Model, DOM）之外；
- 一次编写多次使用：类比的，React 仅作为传统的“模型-视图-控制器”（Model, View, Controller）架构中的视图层，并不对所处系统的技术栈做出假设。因此，用户可以在不重写现有代码的前提下，通过 React 开发新功能。同时，React 也可以和 Node.js 结合支持服务器端渲染，或通过 React Native 赋能移动应用程序。

##### 4.4.2.2 D3

D3 作为 JavaScript 第三方库，主要用于进行在 Web 浏览器中的交互式数据可视化编程，历经多个版本的改动，目前最为主流的为 D3 v4 版本（第四版本）。它遵循主流的 SVG, HTML5 和 CSS 标准，并支持对视觉效果的高度定制化编程。截至今日，已有数以万计的网站使用它作为构建可视分析系统最常用的工具，其作者 Mick Bostock 的画廊[54] 展示了诸多使用 D3 构建的样例。同时，许多鼎鼎有名的第三方开源库也直接使用或借鉴了 D3 的代码或设计理念，这些软件包括但不限于：百度的 ECharts、蚂蚁金服的 G2、Vega 及其拓展 Vega-lite。本文系统遵循业界最佳实践，采用 D3 实现了电网拓扑图、统计图、WaveLines 视图及半矩阵视图。

#### 4.4.3 状态管理器

Redux 是为 JavaScript 应用开发的状态容器。它非常轻量，连同其依赖库，仅有 2KB 大小。它可以帮助开发者编写在不同环境（客户端、服务器和本地）中运行一致并且易于测试的应用程序。例如，它通过支持“结合时间旅行（time travelling）调试器”的实时代码编辑器提供了很好的开发者体验。

根据 Redux 的设计理念，本文作者定义了全局状态用来描述当前应用所包含的服务器响应、缓存数据和尚未发往服务器的局部数据，将其与 React 一起配合使用，构建了复杂的单页面应用。

##### 4.4.4 均衡化算法的优化

算法 3 所描述的均衡化算法可以使原本不均匀分布的数值分布均匀。假设样本中的数

值个数为  $N$ , 那么算法 3 的时间复杂度是为  $O(N^2)$ . 算法 8 运用快速排序及二分查找, 可以将算法的平均时间复杂度优化为  $O(N \log N)$ .

#### 4.5 本章小结

本章主要介绍了本文系统中四个模块（暂态稳定计算模块、数据预处理模块、后端服务器及前端浏览器）各自承担的任务, 并描述了本文作者在系统设计与实现的过程中, 对后三个模块所作出的贡献。

## 第 5 章案例调研

## 第 6 章专家意见

本文作者对一批领域专家进行了采访以验证本文系统的有效性。其中两人是电网模拟计算和稳定性分析的专业分析师, 并参与了本文系统的开发。他们比较熟悉本文系统的视觉设计和交互。其他人则是经验丰富的操作员, 负责模拟数据管理或电网指示。这是他们第一次使用我们的系统来进行可视分析。采访前, 本文作者向受访者详细介绍和说明了本系统的使用方法。每次面试持续了大约 45 分钟。访谈期间, 本文作者向受访者展示了三个典型的数据样本, 并要求他们通过探索系统来分析这些样本。之后, 本文作者问了一些如表 1 所示的关于系统可用性、视觉设计和交互的问题, 并记录了受访者们的反馈意见。

采访后, 被访者普遍赞赏本系统的可视设计与交互, 并认为本系统易用并容易理解。一位分析师评价到: “WaveLiens 和半矩阵的视图都足够明确, 即使没有可视化经验的用户也能够流利地使用它。” 其他评论包括: “不同视图之间的关联交互为用户从各种视图中获取信息提供了便利。而且, 通过时间轴上的缩放操作, 用户可以进一步识别变化图案的细节。” 总的来说, 所有受访者都赞同本系统的可用性。他们也证实, 本系统构建了一个直观、系统的方法, 用于电网的稳定性分析。他们评论说: “WaveLines 是电网分析中多个方面的有效结合”、“WaveLines 提供了一种直观的方法来衡量系统的稳定性, 便于深入分析, 如不稳定性推理和模式识别”和“使用 WaveLines 的用户交互界面来研究不同变量之间的关联关系十分方便”。他们还指出, WaveLines 在电网分析方面具有巨大的潜力。一位分析师说: “我们现在处于一个有意义的起点。我们所做的工作颇具启发性, 因为我们正在为电网分析开辟一个全新的方向。电压和频率的交互相关分析是相当有价值的, 但以前从来没有任何实用工具实现过这一点。”

## 第 7 章总结与展望

### 7.1 全文总结

本文介绍了一种针对电力行业基于仿真的稳定性分析应用领域的可视化分析方法。本文提出了一个多变量模拟时间序列相关分析的可视分析系统, 旨在帮助分析电网的稳定性。特别的, 本文发明了新颖的可视化设计设计和基于统计质量控制的方法在支持异常检测和模式识别中的应用。本文作者与领域专家紧密合作, 提取了一系列分析任务、要求和设计指南来指导系统设计。案例研究和专家反馈证实了本系统的有效性和实用性。就本文作者所知, 本文是国内首次尝试将可视化分析方法应用于电网稳定性分析的工作, 并相信本文所提出的工作有希望激发未来的电网分析, 为构建智能电网添砖加瓦。

### 7.2 未来展望

电网系统的稳定性关系到国民安危与经济增长, 本文作者在设计并实现可视分析系统的

过程中，也发掘了诸多可进一步探究的课题：（1）WaveLines 视图中蕴含的特征为专家提供了从全局到局部的可视分析解决方案，运用深度学习技术，从海量图片中学习出模型，将有可能进一步帮助专家研判电网系统稳定性。（2）相比针对小算例数据的处理方式，本文系统在处理大算例数据时显得相对粗糙：揭示了稳定电网中存在的局部扰动，却没能很好地刻画全网状态。探究针对大规模电网的可视分析技术具有着极其重要的实际意义。（3）基于拼接屏的超高清可视分析技术在国内外学术界和工业界中均有应用，本文作者所在小组正维护着一个由 33 块 55 英寸显示屏所组成的显示集群，将软件重构为适配该硬件的可视分析系统可以带来令人震撼的效果，同时也可超高清拼接屏可视分析的研究积累宝贵经验